



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 137481**

**TO: David Lukton**  
**Location: REM-3B75/3C70**  
**Art Unit: 1653**  
**Tuesday, November 16, 2004**

**Case Serial Number: 09/943002**

**From: Mary Jane Ruhl**  
**Location: Biotech-Chem Library**  
**Remsen 1-A-62**  
**Phone: 571-272-2524**

**[maryjane.ruhl@uspto.gov](mailto:maryjane.ruhl@uspto.gov)**

### **Search Notes**

Examiner Lukton,

Here are the results for your recent search request.

Please feel free to contact me if you have any questions about these results.

Thank you for using STIC services. We appreciate the opportunity to serve you.

Sincerely,

Mary Jane Ruhl  
Technical Information Specialist  
STIC  
Remsen 1-A-62  
Ext. 22524

SEARCH REQUEST FORM  
(STIC)

Access DB# \_\_\_\_\_

Requestor's Name: David Lukton

Examiner number: 71263

Date: 11/9/04

Art Unit: 1653

Phone number: 571-272-0952

Serial Number:

09-943002

Mail Box: 3-C-70

Examiner Rm: 3-B-75

Results format: paper

\*\*\*\*\*

Title: Novel reovirus-derived proteins and uses therefor

Applicant: Roy Duncan

Earliest Priority Date: 11/7/97

~~~~~

Please search SEQ ID NO: 14

RECEIVED  
NOV - 9 2004  
TECH/STIC



```
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21522
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21522

Query Match      11.2%; Score 81; DB 4; Length 271;
Best Local Similarity 26.2%; Pred. No. 0.44;
Matches 32; Conservative 25; Mismatches 53; Indels 12; Gaps 4;

QY 13 PPNAFVEIVSSSTGIIIVGIFAFIPFLYKLLQWNRKSKNKKKEQIREQIELGLLS 72
DB 23 PPRPFPORLHSTGTMTSNDVDAEIAKFEALAHRWDRSEBPKPLHD--INFLRVNWD 80
QY 73 YGAGVASLPLLVIAHNPQSGVISATPIYKGP-CTGVPN-----GRLLQITSGTAENT 124
DB 81 ERAGLAGKVLDIGC--GGGILSEMAQORGANVTGIDMGEPALAVARLHQLSGVAVDYR 138
QY 125 RI 126
DB 139 QI 140

RESULT 3
US-09-710-279-1406
; Sequence 1406, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUB3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1406
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (415)
; OTHER INFORMATION: variable amino acid
US-09-710-279-1406

Query Match      11.2%; Score 80.5; DB 4; Length 415;
Best Local Similarity 34.0%; Pred. No. 0.91;
Matches 17; Conservative 13; Mismatches 19; Indels 1; Gaps 1;

QY 33 GIFAFIPFLYKLLQWNRKSKNKKKEQIREQIELGLLSYGAGV-ASLP 81
DB 159 GVHAEPTEFGVKWALWYTEMKRNLFKEVKEIEVGNKSGAVGTFANIP 208

RESULT 4
US-09-134-001C-4663
; Sequence 4663, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
```

```
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4663
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4663

Query Match      11.2%; Score 80.5; DB 3; Length 458;
Best Local Similarity 34.0%; Pred. No. 1.1;
Matches 17; Conservative 13; Mismatches 19; Indels 1; Gaps 1;

QY 33 GIFAFIPFLYKLLQWNRKSKNKKKEQIREQIELGLLSYGAGV-ASLP 81
DB 169 GVHAEPTEFGVKWALWYTEMKRNLFKEVKEIEVGNKSGAVGTFANIP 218

RESULT 5
US-08-843-530B-36
; Sequence 36, Application US/08843530B
; Patent No. 5939306
; GENERAL INFORMATION:
; APPLICANT: Selitrennikoff, Claude
; APPLICANT: Agnan, Jacqueline
; APPLICANT: Alex, Lisa A.
; APPLICANT: Simon, Melvin I.
; TITLE OF INVENTION: Osmosensing Histidine Kinases
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/843,530B
; FILING DATE: 16-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Macknight, Kamrin T.
; REGISTRATION NUMBER: 38,230
; REFERENCE/DOCKET NUMBER: UTC-02717
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1220 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
US-08-843-530B-36

Query Match      10.5%; Score 75.5; DB 2; Length 1220;
Best Local Similarity 23.3%; Pred. No. 16;
Matches 31; Conservative 22; Mismatches 53; Indels 27; Gaps 6;

QY 1 MCQRHSIVQPPAPPNPAFVEIVSSGTGIIIVGIFAFIPFLYKLLQWNRKSKNKKKE 60
DB 317 VSQPSVFLSPA-----TKLAKITGTVAIGVFVILLTL--PLAHW-----AVQP 360
```



```

; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 26748
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-26748

Query Match          10.1%; Score 72.5; DB 4; Length 338;
Best Local Similarity 22.8%; Pred.No.5.8;
Matches      26; Conservative 17; Mismatches 42; Indels 29; Gaps 3;

QY    34 IPAFSPFYKLLWNRKSNNKKKEQ-----IREQ-----IELGLLSVGAG 76
      |||   ::   |   ::   ||   ||::   ||::   ||:   ||:
Db    158 IFAYIGCVLFEWATWIANKEYDENKEETNLAIRKEWAISLAWYLAPSAALFLLAYGAA 217
      |||   ::   |   ::   ||   ||::   ||::   ||:   ||:

QY    77 VASLPPLNVIAHNPGSVI-----SATPIYKGCTGVPSNRLQLITSG 118
      |||   ::   |   ::   ||   ||::   ||::   ||:   ||:
Db    218 LPFTKLVRILWIPLGPVEVHIKTSYPFIPGRTPVTVPLENLKRKHTRSVWTG 271
      |||   ::   |   ::   ||   ||::   ||::   ||:   ||:

RESULT 12
US-08-843-530B-32
; Sequence 32, Application US/08843530B
; Patent No. 5939306
; GENERAL INFORMATION:
; APPLICANT: Selitrennikoff, Claude
; APPLICANT: Agnan, Jacqueline
; APPLICANT: Alex, Lisa A.
; APPLICANT: Simon, Melvin I.
; TITLE OF INVENTION: Osmosensing Histidine Kinases
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Medlen & Cartoll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/843,530B
; FILING DATE: 16-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MacKnight, Kamrin T.
; REGISTRATION NUMBER: 38,230
; REFERENCE/DOCKET NUMBER: UTC-02717
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 751 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein

```



```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/195,524
; APPLICATION NUMBER: US/09/195,524
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/611,729
; FILING DATE: 06-MAR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-037
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1257 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-09-195-524-8

```

```

Query Match      9.9%; Score 71.5; DB 4; Length 1257;
Best Local Similarity 27.3%; Pred. No. 49;
Matches 42; Conservative 18; Mismatches 63; Indels 31; Gaps 8;

QY      2 GQRHSIVQPPAPPNAP-----VEIV---SSSTGIIIV--GIFAFIFGFLYKLLQ 47
DB      874 GAHAIVAAITQRGNSLLLAIVTEVKYETVTGSGSTGLLVPLCGAFSVLWLCVVLV 933

QY      48 WYNRKSKNKKRQIQIEQLGLLSYGAGVASLPLNLVIAHNPQ-----SVISATPIY 100
DB      934 WYTR---KRRKERERSLPREESANNQWAPLNPINPI-ERPGHKDVLVYQCKNFTPPP 988

QY      101 KGECTGVPSNELLQITSGTAENTRIILNHDGRNP 134
DB      989 RRCPCGRPATR----PSGRM-RRTRILAAVRTP 1017

```

Search completed: November 15, 2004, 14:59:04  
Job time : 42 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 15, 2004, 14:31:00 ; Search time 155 Seconds  
(without alignments)  
324.014 Million cell updates/sec

Title: US-09-943-002-14  
Perfect score: 720  
Sequence: 1 MGQRHSIVQPPAPPNAFV.....EENTRLNHDGRNPDGINSV 140

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : A\_Genesec\_23Sep04:\*
- 1: Genesecqp1980s:\*
  - 2: Genesecqp1990s:\*
  - 3: Genesecqp2000s:\*
  - 4: Genesecqp2001s:\*
  - 5: Genesecqp2002s:\*
  - 6: Genesecqp2003as:\*
  - 7: Genesecqp2003bs:\*
  - 8: Genesecqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description        |
|------------|-------|-------------|--------|-------|--------------------|
| 1          | 720   | 100.0       | 140    | 2     | AAY06113 Baboon re |
| 2          | 84    | 11.7        | 431    | 4     | AAY06113 Baboon re |
| 3          | 84    | 11.7        | 431    | 4     | AAY06113 Baboon re |
| 4          | 84    | 11.7        | 431    | 6     | ABU16411 Protein e |
| 5          | 84    | 11.7        | 454    | 6     | ABM71462 Staphyloc |
| 6          | 84    | 11.7        | 454    | 8     | ADO25483 S aureus  |
| 7          | 81    | 11.2        | 271    | 7     | ABO72776 Pseudomon |
| 8          | 80.5  | 11.2        | 415    | 4     | AAG82156 S. epider |
| 9          | 80.5  | 11.2        | 431    | 6     | ABU43432 Protein e |
| 10         | 80.5  | 11.2        | 431    | 6     | ABU42716 Protein e |
| 11         | 80.5  | 11.2        | 458    | 5     | ABP39818 Staphyloc |
| 12         | 79    | 11.0        | 454    | 8     | ADO25485 S aureus  |
| 13         | 78    | 10.8        | 391    | 5     | ABB48477 Listeria  |
| 14         | 78    | 10.8        | 391    | 6     | ABU32700 Protein e |
| 15         | 77.5  | 10.8        | 339    | 6     | ABU26499 Protein e |
| 16         | 77    | 10.7        | 852    | 4     | ABG13686 Novel hum |
| 17         | 76.5  | 10.6        | 681    | 5     | ABB93650 Herbicida |
| 18         | 75.5  | 10.5        | 427    | 7     | ADK61938 Disease t |
| 19         | 75.5  | 10.5        | 1220   | 8     | ADN07100 S. cerevi |
| 20         | 75    | 10.4        | 892    | 7     | ADJ68848 Human hea |
| 21         | 74.5  | 10.3        | 231    | 6     | ABR41590 Human DIT |
| 22         | 74.5  | 10.3        | 319    | 6     | ABM72920 Staphyloc |
| 23         | 74.5  | 10.3        | 498    | 4     | AAG92924 C glutami |
| 24         | 74    | 10.3        | 408    | 2     | AAW75856 Human sec |
| 25         | 74    | 10.3        | 408    | 2     | AAY45161 Human sec |

|    |      |      |      |   |                    |
|----|------|------|------|---|--------------------|
| 26 | 74   | 10.3 | 1455 | 4 | AAM80143 Human pro |
| 27 | 74   | 10.3 | 1726 | 3 | AAY57572 Moraxella |
| 28 | 74   | 10.3 | 1739 | 8 | ADL06053 M. catarr |
| 29 | 73.5 | 10.2 | 348  | 6 | ADA48468 Rice prot |
| 30 | 73.5 | 10.2 | 494  | 4 | ABE61632 Drosophil |
| 31 | 73.5 | 10.2 | 534  | 2 | AAY26967 Wheat Mlo |
| 32 | 73.5 | 10.2 | 534  | 3 | AAH01805 Wheat Mlo |
| 33 | 73.5 | 10.2 | 534  | 4 | AAH31251 Amino aci |
| 34 | 73   | 10.1 | 262  | 6 | ABU31472 Protein e |
| 35 | 73   | 10.1 | 482  | 6 | ADA36497 Acinetoba |
| 36 | 73   | 10.1 | 543  | 8 | ADH71860 Human pro |
| 37 | 73   | 10.1 | 766  | 5 | ABG96273 Human imm |
| 38 | 73   | 10.1 | 766  | 8 | ADH71856 Human pro |
| 39 | 72.5 | 10.1 | 534  | 2 | AAY26968 Wheat Mlo |
| 40 | 72.5 | 10.1 | 534  | 4 | AAH31252 Amino aci |
| 41 | 72.5 | 10.1 | 751  | 8 | ADN07097 S. cerevi |
| 42 | 72.5 | 10.1 | 1677 | 8 | ADP76092 Arabidops |
| 43 | 72   | 10.0 | 649  | 8 | ADM48283 Polypepti |
| 44 | 72   | 10.0 | 687  | 7 | ADB64828 Human pro |
| 45 | 72   | 10.0 | 775  | 4 | AAB93302 Human pro |

ALIGNMENTS

RESULT 1  
ID AAY06113 standard; protein; 140 AA.  
XX AC AAY06113;  
DT 16-AUG-1999 (first entry)  
DE Baboon reovirus fusogenic protein P15a.  
XX KW Orthoreovirus; reovirus; BRV; P15 protein; fusogenic protein;  
cell fusion; membrane fusion; syncytium formation.  
XX OS Baboon reovirus.  
XX FH Key Location/Qualifiers  
FT Domain 27..41  
FT /note= "transmembrane domain"  
XX PN WO9924582-A1.  
XX PD 20-MAY-1999.  
XX PF 06-NOV-1998; 98WO-CA001046.  
XX PR 07-NOV-1997; 97US-00965708.  
XX PA (UYDA-) UNIV DALHOUSIE.  
XX PI Duncan R;  
XX DR WPI; 1999-327410/27.  
XX DR N-PSDB; AAX58670.  
XX PT Fusogenic proteins from non-enveloped reoviruses.  
XX PS Example 6; Page 53-54; 57pp; English.  
XX CC The present sequence represents the fusogenic protein P15a of baboon reovirus (BRV). 2 Unrelated fusion proteins responsible for cell-cell fusion induced by avian reovirus (ARV) and the only 2 fusogenic mammalian reoviruses, Nelson Bay virus (NBV) and BRV, are identified in this invention. These proteins are termed P11 for ARV and NBV (see also AAY06104, AAY06107 and AAY06110) and P15 for BRV. Fusogenic P11 and P15 are useful for (i) promoting fusion of cell, liposome or proteoliposome membranes; production of B and T cell hybridomas or other heterokaryons (they eliminate the need to use toxic chemicals such as PEG); (iii) for producing fusions between liposomes and cells or other liposomes; and

CC (iv) for incorporation into liposomes to allow intra- or extracellular  
 CC delivery of bioactive agents (e.g. nucleic acid, proteins and metabolic  
 CC regulators), both in vivo and in cell cultures. Antibodies raised against  
 CC P11 or P15 are useful for diagnosis and therapeutically (e.g. to block  
 CC undesirable fusion processes). P11 and P15 are smaller than fusogenic  
 CC proteins from enveloped viruses and are not glycosylated, so are easier  
 CC to produce and purify using recombinant expression systems. They do not  
 CC require a signal peptide for membrane entry, are not immunogenic and are  
 CC functional at neutral pH

XX Sequence 140 AA;

4 Query Match 100.0%; Score 720; DB 2; Length 140;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-72;  
 Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQRHSIVOPAPPNAPFVIVSSSTGIIIVGIFAFIFSLYKLLQWYNRKSKNKKKE 60

DB 1 MQRHSIVOPAPPNAPFVIVSSSTGIIIVGIFAFIFSLYKLLQWYNRKSKNKKKE 60

QY 61 QIREQIELGLLYGAGVASLPLNLVIAHNPVSGVISATPIYKGPCTGVPSRLLOITSGTA 120

DB 61 QIREQIELGLLYGAGVASLPLNLVIAHNPVSGVISATPIYKGPCTGVPSRLLOITSGTA 120

QY 121 ENTRILNHDGRNPDGSINV 140

DB 121 ENTRILNHDGRNPDGSINV 140

#### RESULT 2

AAU33910  
 ID AAU33910 standard; protein; 431 AA.

XX AAU33910;

DT 14-FEB-2002 (first entry)

XX Staphylococcus aureus cellular proliferation protein #186.

DE Antisense; prokaryotic cellular proliferation protein; antibiotic;  
 KW antibacterial; drug design.

XX Staphylococcus aureus.

OS WO200170955-A2.

PN 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US009180.

XX 21-MAR-2000; 2000US-0191078P.

PR 23-MAY-2000; 2000US-0206848P.

PR 26-MAY-2000; 2000US-0207727P.

PR 23-OCT-2000; 2000US-0242578P.

PR 27-NOV-2000; 2000US-0253625P.

PR 22-DEC-2000; 2000US-0257931P.

PR 16-FEB-2001; 2001US-0269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GU;

PI Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

DR N-PSDB; AAS51769.

XX New polynucleotides for the identification and development of  
 PT antibiotics, comprise sequences of antisense nucleic acids.

XX Example 3; SEQ ID NO 5406; 51pp; English.

XX The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the genes,

CC their use in the discovery of novel antibiotics, the essential genes  
 CC themselves and the encoded proteins. The prokaryotes used are Escherichia  
 CC coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,  
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also  
 CC useful for the identification of potential new targets for antibiotic  
 CC development. The antisense nucleic acids can also be used to identify  
 CC proteins used in proliferation, to express these proteins, and to obtain  
 CC antibodies capable of binding to the expressed proteins. The proteins can  
 CC be used to screen compounds in rational drug discovery programmes. The  
 CC antisense nucleic acid sequence is also useful to screen for homologous  
 CC nucleic acids which are required for cell proliferation in a wide variety  
 CC of organisms. The present sequence represents an essential prokaryotic  
 CC cellular proliferation protein. Note: The sequence data for this patent  
 CC did not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 431 AA;

Query Match 11.7%; Score 84; DB 4; Length 431;  
 Best Local Similarity 26.5%; Pred. No. 2.2;  
 Matches 26; Conservative 15; Mismatches 35; Indels 22; Gaps 2;

QY 33 GIFAFIFSLYKLLQWYNRKSKNKKKEQIREQIELGLLYGAGVASLPLNLVIAHNP 92

DB 142 GVHAEPTTFGVKMLWYTEMQRNLQRPQVREIEVGKMSGAVG-----TFANIPPE 193

QY 93 VISATPIYKGPCTGVPSRLLOITSGTAENTRIILNHD 130

DB 194 IESYVCKHLG-----IGTAPVSTQLQRD 217

#### RESULT 3

AAU36895  
 ID AAU36895 standard; protein; 431 AA.

XX AAU36895;

DT 14-FEB-2002 (first entry)

XX Staphylococcus aureus cellular proliferation protein #1085.

DE Antisense; prokaryotic cellular proliferation protein; antibiotic;  
 KW antibacterial; drug design.

XX Staphylococcus aureus.

OS WO200170955-A2.

PN 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US009180.

XX 21-MAR-2000; 2000US-0191078P.

PR 23-MAY-2000; 2000US-0206848P.

PR 26-MAY-2000; 2000US-0207727P.

PR 23-OCT-2000; 2000US-0242578P.

PR 27-NOV-2000; 2000US-0253625P.

PR 22-DEC-2000; 2000US-0257931P.

PR 16-FEB-2001; 2001US-0269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GU;

PI Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

DR N-PSDB; AAS54754.

XX New polynucleotides for the identification and development of  
 PT antibiotics, comprise sequences of antisense nucleic acids.

XX Example 3; SEQ ID NO 12488; 51pp; English.

XX The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the genes,  
 CC themselves and the discovery of novel antibiotics, the essential genes  
 CC colli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,  
 CC Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also  
 CC useful for the identification of potential new targets for antibiotic  
 CC development. The antisense nucleic acids can also be used to identify  
 CC proteins used in proliferation, to express these proteins, and to obtain  
 CC antibodies capable of binding to the expressed proteins. The proteins can  
 CC be used to screen compounds in rational drug discovery programmes. The  
 CC antisense nucleic acid sequence is also useful to screen for homologous  
 CC nucleic acids which are required for cell proliferation in a wide variety  
 CC of organisms. The present sequence represents an essential prokaryotic  
 CC cellular proliferation protein. Note: The sequence data for this patent  
 CC did not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 431 AA;

Query Match 11.7%; Score 84; DB 4; Length 431;  
 Best Local Similarity 26.5%; Pred. No. 2.2;  
 : Matches 26; Conservative 15; Mismatches 35; Indels 22; Gaps 2;  
 QY 33 GIFAFIFLYKLLQWNRKSKKKEQIREQIETLGLLSYGAGVASLPLLNVAHNPGS 92  
 Db 142 GVHAEPTTGVKQWALWTEQRLQPKQVRREIEVGKMSGAVG-----TFANIPPE 193  
 QY 93 VISATPIYKPGCTGVPSNRLQLTQTSPTAENTRLNHD 130  
 Db 194 IESVCKHLG-----IGTAPVSTQTLQRD 217

RESULT 4  
 ABU16411  
 ID ABU16411 standard; protein; 431 AA.  
 XX  
 AC ABU16411;  
 DT 19-JUN-2003 (first entry)  
 XX Protein encoded by Prokaryotic essential gene #1938.  
 DE Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 KW Staphylococcus aureus.  
 OS WO200277183-A2.  
 PN 03-OCT-2002.  
 PP 21-MAR-2002; 2002WO-US009107.  
 PF 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362699P.  
 XX (ELIT-) ELITRA PHARM INC.  
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
 PI Wall D, Trawick JD, Carr GU, Yamamoto R, Foreyth RA, Xu HH;  
 XX WPI; 2003-029926/02.  
 DR N-PSDB; ACA20281.  
 XX New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX

Claim 25; SEQ ID NO 44335; 1766pp; English.

PS The invention relates to an isolated nucleic acid comprising any one of  
 XX the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 431 AA;

Query Match 11.7%; Score 84; DB 6; Length 431;  
 Best Local Similarity 26.5%; Pred. No. 2.2;  
 Matches 26; Conservative 15; Mismatches 35; Indels 22; Gaps 2;  
 QY 33 GIFAFIFLYKLLQWNRKSKKKEQIREQIETLGLLSYGAGVASLPLLNVAHNPGS 92  
 Db 142 GVHAEPTTGVKQWALWTEQRLQPKQVRREIEVGKMSGAVG-----TFANIPPE 193  
 QY 93 VISATPIYKPGCTGVPSNRLQLTQTSPTAENTRLNHD 130  
 Db 194 IESVCKHLG-----IGTAPVSTQTLQRD 217

RESULT 5  
 ABM71462  
 ID ABM71462 standard; protein; 454 AA.  
 XX  
 AC ABM71462;  
 XX 20-NOV-2003 (first entry)  
 DT Staphylococcus aureus protein #702.  
 DE  
 DE Staphylococcus aureus protein #702.  
 XX Antibacterial; vaccine; gene therapy; infection; sepsis; diagnosis;  
 KW enzymatic assay; antibiotic target.  
 KW Staphylococcus aureus.  
 OS WO200294868-A2.  
 XX 28-NOV-2002.  
 PD 27-MAR-2002; 2002WO-IB002637.  
 PF 27-MAR-2001; 2001GB-00007661.  
 XX (CHIR-) CHIRON SPA.  
 PA

```

XX Masignani V, Mora M, Scarselli M;
XX WPI: 2003-120786/11.
DR N-PSDB, ACP73022.
XX
PT New Staphylococcus aureus protein, useful as a vaccine for treating or
PT preventing Staphylococcal infection, specifically an infection caused by
PT S. aureus, e.g. sepsis.
XX
PS Claim 1; SEQ ID NO 1404; 49pp; English.
XX
CC The invention relates to novel genes and encoded proteins from
CC Staphylococcus aureus. A composition comprising the S. aureus protein, a
CC nucleic acid encoding the protein, or an antibody to the protein, is
CC useful as a pharmaceutical, particularly as a vaccine for treating or
CC preventing infection due to Staphylococcus bacteria, specifically an
CC infection caused by S. aureus. The composition is particularly useful for
CC treating or preventing sepsis in a patient. The composition can also be
CC used for diagnostics. The protein is also used in an assay for enzymatic
CC studies and as a target for antibiotics. This sequence represents one of
CC the novel S. aureus proteins of the invention
XX
SQ Sequence 454 AA;
    Query Match      11.7%; Score 84; DB 6; Length 454;
    Best Local Similarity 26.5%; Pred. No. 2.3;
    Matches 26; Conservative 15; Mismatches 35; Indels 22; Gaps 2;
XX
QY 33 GIFAFIPFLYKLLQWNRKSKNQKKEIQRIQELGLLSYGAGVASPLLNVIANNPGS 92
Db 165 GVHAEPTTFGVKALWYTEMQNLORFKQVREIEVGMKGAVG-----TFANIPPE 216
QY 93 VISATPIYKGPCTGVNSRLLOITSGTAENTRIINHND 130
Db 217 IBSYVCKHLG-----IGTAPVSTQLORD 240
XX
RESULT 6
ID ADO25483
XX ADO25483 standard; protein; 454 AA.
XX
AC ADO25483;
XX
BT 12-AUG-2004 (first entry)
XX
DE S_aureus adenylosuccinate lyase purB protein Seq158.
XX
KW antimicrobial; microbial disease; drug composition; vaccine;
KW bacterial infection; antibacterial; food preservative.
XX
OS Staphylococcus aureus.
XX
PN WO2004041854-A2.
XX
PD 21-MAY-2004.
XX
PP 05-NOV-2003; 2003WO-CA001671.
XX
PR 05-NOV-2002; 2002US-0423757P.
PR 05-NOV-2002; 2002US-0423758P.
PR 05-NOV-2002; 2002US-0423791P.
PR 05-NOV-2002; 2002US-0423832P.
PR 05-NOV-2002; 2002US-0423875P.
PR 05-NOV-2002; 2002US-0423915P.
PR 06-NOV-2002; 2002US-0424362P.
PR 06-NOV-2002; 2002US-0424367P.
PR 06-NOV-2002; 2002US-0424370P.
PR 06-NOV-2002; 2002US-0424373P.
PR 06-NOV-2002; 2002US-0424376P.
PR 06-NOV-2002; 2002US-0424389P.
PR 07-NOV-2002; 2002US-0424502P.
PR 07-NOV-2002; 2002US-0424651P.
XX
QY 07-NOV-2002; 2002US-0424664P.
PR 07-NOV-2002; 2002US-0424665P.
PR 08-NOV-2002; 2002US-0424968P.
PR 08-NOV-2002; 2002US-0425076P.
PR 08-NOV-2002; 2002US-0425085P.
PR 08-NOV-2002; 2002US-0425118P.
PR 08-NOV-2002; 2002US-0425128P.
PR 08-NOV-2002; 2002US-0425162P.
PR 08-NOV-2002; 2002US-0425201P.
PR 12-MAR-2003; 2003US-0453914P.
PR 12-MAR-2003; 2003US-0454021P.
PR 12-MAR-2003; 2003US-0454128P.
PR 12-MAR-2003; 2003US-0454193P.
PR 13-MAR-2003; 2003US-0454215P.
PR 13-MAR-2003; 2003US-0454218P.
PR 13-MAR-2003; 2003US-0454487P.
PR 13-MAR-2003; 2003US-0454507P.
PR 13-MAR-2003; 2003US-0454538P.
PR 14-MAR-2003; 2003US-0455010P.
PR 14-MAR-2003; 2003US-0455036P.
PR 14-MAR-2003; 2003US-0455054P.
PR 14-MAR-2003; 2003US-0455082P.
PR 17-MAR-2003; 2003US-0455191P.
PR 17-MAR-2003; 2003US-0455192P.
PR 17-MAR-2003; 2003US-0455334P.
PR 17-MAR-2003; 2003US-0455335P.
PR 17-MAR-2003; 2003US-0455343P.
XX
XX (AFPI-) AFFINIUM PHARM INC.
XX
PI Edwards A, Dharamsi A, Vedadi M, Domagala M, Nethery K;
PI Mansoury K, Pinder B, Alam MZ, Ng I, Virag C, Houston S;
PI McDonald M, Buzadzija K;
XX
DR WPI: 2004-400642/37.
DR N-PSDB; ADO25482.
XX
XX Bacterial polypeptide composition useful for treating bacterial
PT infection, has isolated, recombinant bacterial polypeptide such as GTP-
PT binding protein Bra from Pseudomonas aeruginosa or adenylosuccinate lyase
PT from Enterococcus faecalis.
XX
PS Claim 41; SEQ ID NO 158; 566pp; English.
XX
CC This invention relates to the identification of novel protein targets for
CC the development of antimicrobial drugs against pathogenic bacteria.
CC Specifically, it refers to recombinant proteins derived from
CC Staphylococcus aureus, Helicobacter pylori, Streptococcus pneumoniae,
CC Escherichia coli, Enterococcus faecalis or Pseudomonas aeruginosa. The
CC present invention describes providing a three-dimensional structure for
CC these crystallised proteins to identify a potential modulator for the
CC prevention or treatment of microbial diseases. Furthermore, contacting a
CC protein with a modulator can be useful for assaying protein activity and
CC hence its viability in drug composition or vaccine. Accordingly, such
CC compositions can be useful for treating bacterial infections, developing
CC antibacterial agents useful as food preservatives or treating food
CC products to eliminate potential pathogens. This polypeptide sequence is a
CC bacterial protein target of the invention.
XX
SQ Sequence 454 AA;
    Query Match      11.7%; Score 84; DB 8; Length 454;
    Best Local Similarity 26.5%; Pred. No. 2.3;
    Matches 26; Conservative 15; Mismatches 35; Indels 22; Gaps 2;
XX
QY 33 GIFAFIPFLYKLLQWNRKSKNQKKEIQRIQELGLLSYGAGVASPLLNVIANNPGS 92
Db 165 GVHAEPTTFGVKALWYTEMQNLORFKQVREIEVGMKGAVG-----TFANIPPE 216
QY 93 VISATPIYKGPCTGVNSRLLOITSGTAENTRIINHND 130
Db 217 IBSYVCKHLG-----IGTAPVSTQLORD 240
XX

```



19-JUN-2003 (first entry)  
 Protein encoded by Prokaryotic essential gene #28959.  
 Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 Staphylococcus haemolyticus.  
 WO200277183-A2.  
 03-OCT-2002.  
 21-MAR-2002; 2002WO-US009107.  
 21-MAR-2001; 2001US-00815242.  
 06-SEP-2001; 2001US-00948993.  
 25-OCT-2001; 2001US-0342923P.  
 08-FEB-2002; 2002US-00072851.  
 06-MAR-2002; 2002US-0362699P.  
 (ELIT-) ELITRA PHARM INC.  
 Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,  
 Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 WPI; 2003-029926/02.  
 N-PSDB; ACA47302.  
 New antisense nucleic acids, useful for identifying proteins or screening  
 for homologous nucleic acids required for cellular proliferation to  
 isolate candidate molecules for rational drug discovery programs.  
 Claim 25; SEQ ID NO 71356; 1766pp; English.  
 The invention relates to an isolated nucleic acid comprising any one of  
 the 6213 antisense sequences given in the specification where expression  
 of the nucleic acid inhibits proliferation of a cell. Also included are:  
 (1) a vector comprising a promoter operably linked to the nucleic acid  
 encoding a polypeptide whose expression is inhibited by the antisense  
 nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 polypeptide or its fragment whose expression is inhibited by the  
 antisense nucleic acid; (4) an antibody capable of specifically binding  
 the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 proliferation or the activity of a gene in an operon required for  
 proliferation; (7) identifying a compound that influences the activity of  
 the gene product or that has an activity against a biological pathway  
 required for proliferation, or that inhibits cellular proliferation; (8)  
 identifying a gene required for cellular proliferation or the biological  
 pathway in which a proliferation-required gene or its gene product lies  
 or a gene on which the test compound that inhibits proliferation of an  
 organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 compound's activity; (11) a culture comprising strains in which the gene  
 product is overexpressed or underexpressed; (12) determining the extent  
 to which each of the strains is present in a culture or collection of  
 strains; or (13) identifying the target of a compound that inhibits the  
 proliferation of an organism. The antisense nucleic acids are useful for  
 identifying proteins or screening for homologous nucleic acids  
 required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
*K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 the target prokaryotic essential genes. Note: The sequence data for this  
 patent did not form part of the printed specification, but was obtained  
 from ftp.wipo.int/pub/published\_pcr\_sequences

Query Match  
 Best Local Similarity 11.2%; Score 80.5; DB 6; Length 431;  
 Matches 17; Conservative 13; Mismatches 19; Indels 1; Gaps 1;  
 33 GIFAFIFSLYKLLQWYRKNKRRKEQIREQIELGLLSYGAGV-ASLP 81

Db 142 GVHAEPFTFGVQVALWYTEMKGNLKRFEVRKEIEVGKMGAVGTANIP 191  
 RESULT 10  
 ABU42716  
 ID ABU42716 standard; protein; 431 AA.  
 XX AC ABU42716;  
 XX DT 19-JUN-2003 (first entry)  
 XX DE Protein encoded by Prokaryotic essential gene #28243.  
 XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
 XX OS Staphylococcus epidermidis.  
 XX PN WO200277183-A2.  
 XX PD 03-OCT-2002.  
 XX PF 21-MAR-2002; 2002WO-US009107.  
 XX PR 21-MAR-2001; 2001US-00815242.  
 XX PR 06-SEP-2001; 2001US-00948993.  
 XX PR 25-OCT-2001; 2001US-0342923P.  
 XX PR 08-FEB-2002; 2002US-00072851.  
 XX PR 06-MAR-2002; 2002US-0362699P.  
 XX PA (ELIT-) ELITRA PHARM INC.  
 XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,  
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
 XX WPI; 2003-029926/02.  
 XX DR N-PSDB; ACA46586.  
 XX PT New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 XX PS Claim 25; SEQ ID NO 70640; 1766pp; English.  
 XX CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 6213 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antisense nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
 CC the target prokaryotic essential genes. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pcr\_sequences



PT binding protein Bra from Pseudomonas aeruginosa or adenylosuccinate lyase  
PT from Enterococcus faecalis.

PS Claim 41; SEQ ID NO 160; 566pp; English.

XX This invention relates to the identification of novel protein targets for  
CC the development of antimicrobial drugs against pathogenic bacteria.  
CC Specifically, it refers to recombinant proteins derived from  
CC Staphylococcus aureus, Helicobacter pylori, Streptococcus pneumoniae,  
CC Escherichia coli, Enterococcus faecalis or Pseudomonas aeruginosa. The  
CC present invention describes providing a three-dimensional structure for  
CC these crystallised proteins to identify a potential modulator for the  
CC prevention or treatment of microbial diseases. Furthermore, contacting a  
CC protein with a modulator can be useful for assaying protein activity and  
CC hence its viability in drug composition or vaccine. Accordingly, such  
CC compositions can be useful for treating bacterial infections, developing  
CC antibacterial agents useful as food preservatives or treating food  
CC products to eliminate potential pathogens. This polypeptide sequence is a  
CC bacterial protein target of the invention.

XX Sequence 454 AA;

Query Match 11.0%; Score 79; DB 8; Length 454;

Best Local Similarity 25.5%; Pred. No. 8.6;

Matches 25; Conservative 16; Mismatches 35; Indels 22; Gaps 2;

QY 33 GIPAFIFSLYKLOWNRKSKKKKEQIREQIELGLLSYGAVASPLLNVAHNPQS 92

DB 165 GVHAPFTTGVQWALYTEMPRNLQRFKQREIEVGRMGAVG-----TFANIPPE 216

QY 93 VISATPIYKQCTGVNSRLLOITSGTAEENTRLNHD 130

DB 217 IESYVCKHLG-----IGTAPVSIQTQLORD 240

RESULT 13

ABB48477

ID ABB48477 standard; protein; 391 AA.

XX ABB48477;

DT 05-FEB-2002 (first entry)

XX Listeria monocytogenes protein #1181.

XX Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;  
KW vitamin B12; bacterial infection; disease.

XX Listeria monocytogenes.

XX WO200177335-A2.

XX 18-OCT-2001.

XX 11-APR-2001; 2001WO-FR001118.

XX 11-APR-2000; 2000FR-00004629.

XX (INSP ) INST PASTEUR.

XX Buchrieser C, Frangeul L, Couve E, Ruenick C, Fsihi H, Dehoux P;  
PI Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;  
PI Daniels J, Goebel W, Kref J, Kuhn M, Ng E, Vazquez-Boland JA;  
PI Dominguez-Bernal G, Garrido-Garcia P, Rierrez-Martinez A, Amend A;  
PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durand L;  
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;  
PI Maduenio E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;  
PI Rose M, Voss H;

XX WPI; 2002-010914/01.

XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment  
PT and prevention of Listeria and related bacterial infections, and related

PT polypeptides.

PS Claim 6; SEQ ID NO 1182; 192pp; French.

XX The present invention relates to the genome sequence of Listeria  
CC monocytogenes EGD-a (see ABA03041). The genome sequence and fragments of  
CC it are useful for selecting probes and primers for detecting genes in L.  
CC monocytogenes and related organisms, and for studying genetic  
CC polymorphisms and other genomes. The present sequence is a protein  
CC encoded by the genome sequence of the present invention. Proteins  
CC expressed from the genome sequence are useful for raising specific  
CC antibodies, identification of L. monocytogenes and related organisms, and  
CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin  
CC B12. The genome sequence and proteins encoded by it are also useful for  
CC selecting compounds that regulate gene expression and cell replication  
CC and modulate L. monocytogenes-related diseases. In addition, the genome  
CC sequence and proteins encoded by it are useful in pharmaceutical and  
CC vaccine compositions for the treatment or prevention of infections by L.  
CC monocytogenes and related organisms. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 391 AA;

Query Match 10.8%; Score 78; DB 5; Length 391;

Best Local Similarity 28.3%; Pred. No. 9.1;

Matches 30; Conservative 14; Mismatches 24; Indels 38; Gaps 5;

QY 56 KKKKQIREQIELGL-----LSYGAVASPLLNVAHNPQS-----VISATPIY 100

DB 79 KKAABESREQIREALKGSDMVFVTAGMGGGTGGAAPVIAQIAKENGALTGVVTRPFGF 138

QY 101 KGP-----CTG-----VPNSRLQITSGTAEENTRIL 127

DB 139 EGPKTKQALGTAEAKGAVDTLIVIPDRLLQI-----VDKNTFML 180

RESULT 14

ABU32700

ID ABU32700 standard; protein; 391 AA.

XX ABU32700;

XX 19-JUN-2003 (first entry)

XX Protein encoded by prokaryotic essential gene #18227.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Listeria monocytogenes.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX N-PSDB; ACA36570.

XX New antisense nucleic acids, useful for identifying proteins or screening



PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
PS Claim 25; SEQ ID NO 60624; 1766pp; English.  
XX  
CC The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 391 AA;  
Query Match 10.8%; Score 78; DB 6; Length 391;  
Best Local Similarity 28.3%; Pred. No. 9.1; Indels 24; Gaps 5;  
Matches 30; Conservative 14; Mismatches 38; Gaps 5;  
QY 56 KKRQIRQIRQLGL-----LSYGAGVASLPLNVAHNPQS----VISATPIY 100  
DB 79 KKAERSRQIRSEALKSDWFTACMGCGGTGTAAPVIAQIAKENGALTGVVTRPFGF 138  
QY 101 KGP-----CTG-----VPSRLQLQITSGTAEENTRIL 127  
DB 139 EGPKRTKQALGTGAEMKAVDTLIVPNDRLQLI-----VDKNTPL 180  
RESULT 15  
ABU26499  
ID ABU26499 standard; protein; 339 AA.  
XX  
AC ABU26499;  
XX  
DT 19-JUN-2003 (first entry)  
XX  
DE Protein encoded by Prokaryotic essential gene #12026.  
XX  
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.  
XX  
OS Campylobacter jejuni.  
XX  
PN WO200277183-A2.  
XX  
PD 03-OCT-2002.  
XX  
PP 21-MAR-2002; 2002WO-US009107.  
XX  
PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.  
PR 25-OCT-2001; 2001US-0342923P.  
PR 08-FEB-2002; 2002US-00072851.  
PR 06-MAR-2002; 2002US-0362699P.  
XX  
PA (ELIT-) ELITRA PHARM INC.  
XX  
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;  
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;  
XX  
DR WPI: 2003-029926/02.  
DR N-PSDB; ACA30369.  
XX  
PT New antisense nucleic acids, useful for identifying proteins or screening  
PT for homologous nucleic acids required for cellular proliferation to  
PT isolate candidate molecules for rational drug discovery programs.  
XX  
PS Claim 25; SEQ ID NO 54423; 1766pp; English.  
XX  
CC The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 339 AA;  
Query Match 10.8%; Score 77.5; DB 6; Length 339;  
Best Local Similarity 28.4%; Pred. No. 8.6;  
Matches 31; Conservative 14; Mismatches 31; Indels 33; Gaps 5;  
QY 48 WYNRKSK-----KKRK-----EQIREQLGLSYGAGVASLPLNVAHNP 91  
DB 61 WLDKEKNILNQKALKNPKLLSIILDKINDELEFGFRFYA-----VKIIAHX 112  
QY 92 SVISATPIYKPGCTGVPSRLQLQITSGTAEENTRILNHDGRNPDGSINV 140  
DB 113 TIVSATDI-----SDEKNERLASMGSAVAH-----LAHEIRNPIGSI 152  
Search completed: November 15, 2004, 14:54:18  
Job time : 164 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 15, 2004, 14:51:55 ; Search time 141 Seconds  
(without alignments)

351.310 Million cell updates/sec

Title: US-09-943-002-14

Perfect score: 720

Sequence: 1 MGQRHSIVQPPAPPNAFVE.....BENTRILNHDGRNPDGSINV 140

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1568699 seqs, 353819137 residues

Total number of hits satisfying chosen parameters: 1568699

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgm2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pap:\*
- 2: /cgm2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pap:\*
- 3: /cgm2\_6/ptodata/2/pubpaa/US05\_NEW\_PUB.pap:\*
- 4: /cgm2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pap:\*
- 5: /cgm2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pap:\*
- 6: /cgm2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pap:\*
- 7: /cgm2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pap:\*
- 8: /cgm2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pap:\*
- 9: /cgm2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pap:\*
- 10: /cgm2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pap:\*
- 11: /cgm2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pap:\*
- 12: /cgm2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pap:\*
- 13: /cgm2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pap:\*
- 14: /cgm2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pap:\*
- 15: /cgm2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pap:\*
- 16: /cgm2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pap:\*
- 17: /cgm2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pap:\*
- 18: /cgm2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pap:\*
- 19: /cgm2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pap:\*
- 20: /cgm2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description          |
|------------|-------|-------------|--------|-------|----------------------|
| 1          | 720   | 100.0       | 140    | 9     | US-09-943-002-14     |
| 2          | 84    | 11.7        | 431    | 9     | US-09-815-242-5406   |
| 3          | 84    | 11.7        | 431    | 9     | US-09-815-242-12488  |
| 4          | 84    | 11.7        | 431    | 15    | US-10-282-122A-44335 |
| 5          | 84    | 11.7        | 366    | 17    | US-10-425-114-45755  |
| 6          | 80.5  | 11.2        | 376    | 17    | US-10-739-930-8926   |
| 7          | 80.5  | 11.2        | 431    | 15    | US-10-282-122A-70640 |
| 8          | 80.5  | 11.2        | 431    | 15    | US-10-282-122A-71356 |
| 9          | 80.5  | 11.2        | 756    | 15    | US-10-424-599-200688 |
| 10         | 80.5  | 11.2        | 760    | 15    | US-10-425-114-39701  |
| 11         | 79.5  | 11.0        | 539    | 17    | US-10-425-115-290480 |
| 12         | 79.5  | 11.0        | 544    | 15    | US-10-425-114-43526  |
| 13         | 78.5  | 10.9        | 419    | 15    | US-10-437-963-183058 |

|    |      |      |      |    |                      |                    |
|----|------|------|------|----|----------------------|--------------------|
| 14 | 78   | 10.8 | 325  | 14 | US-10-369-493-10884  | Sequence 10884, A  |
| 15 | 78   | 10.8 | 391  | 15 | US-10-282-122A-60624 | Sequence 60624, A  |
| 16 | 77.5 | 10.8 | 339  | 15 | US-10-282-122A-54423 | Sequence 54423, A  |
| 17 | 77.5 | 10.8 | 606  | 15 | US-10-425-114-54999  | Sequence 54999, A  |
| 18 | 77.5 | 10.8 | 654  | 14 | US-10-369-493-5059   | Sequence 5059, Ap  |
| 19 | 75.5 | 10.5 | 1220 | 9  | US-09-801-368-332    | Sequence 332, App  |
| 20 | 75.5 | 10.5 | 1220 | 14 | US-10-369-493-22082  | Sequence 22082, A  |
| 21 | 75   | 10.4 | 601  | 16 | US-10-437-963-149924 | Sequence 149924, A |
| 22 | 75   | 10.4 | 892  | 16 | US-10-408-765A-654   | Sequence 654, App  |
| 23 | 74.5 | 10.3 | 456  | 15 | US-10-425-114-53999  | Sequence 53999, A  |
| 24 | 74.5 | 10.3 | 498  | 9  | US-09-738-626-6678   | Sequence 6678, Ap  |
| 25 | 74   | 10.3 | 289  | 15 | US-10-424-599-150408 | Sequence 150408, A |
| 26 | 74   | 10.3 | 468  | 14 | US-10-369-493-4998   | Sequence 4998, Ap  |
| 27 | 73   | 10.1 | 262  | 15 | US-10-282-122A-59396 | Sequence 59396, A  |
| 28 | 73   | 10.1 | 307  | 15 | US-10-425-114-41480  | Sequence 41480, A  |
| 29 | 73   | 10.1 | 359  | 17 | US-10-425-115-302682 | Sequence 302682, A |
| 30 | 73   | 10.1 | 451  | 17 | US-10-425-115-312699 | Sequence 312699, A |
| 31 | 73   | 10.1 | 456  | 15 | US-10-425-114-51670  | Sequence 51670, A  |
| 32 | 73   | 10.1 | 766  | 15 | US-10-471-449-11     | Sequence 11, Appl  |
| 33 | 72   | 10.0 | 163  | 17 | US-10-425-115-274608 | Sequence 274608, A |
| 34 | 72   | 10.0 | 649  | 14 | US-10-310-154-701    | Sequence 701, App  |
| 35 | 72   | 10.0 | 687  | 14 | US-10-104-047-2982   | Sequence 2982, Ap  |
| 36 | 71.5 | 9.9  | 69   | 17 | US-10-425-115-190895 | Sequence 190895, A |
| 37 | 71.5 | 9.9  | 158  | 16 | US-10-437-963-160666 | Sequence 160666, A |
| 38 | 71.5 | 9.9  | 210  | 15 | US-10-282-122A-52165 | Sequence 52165, A  |
| 39 | 71.5 | 9.9  | 241  | 15 | US-10-320-797-3051   | Sequence 3051, Ap  |
| 40 | 71.5 | 9.9  | 481  | 14 | US-10-369-493-20773  | Sequence 20773, A  |
| 41 | 71.5 | 9.9  | 766  | 17 | US-10-487-421-8      | Sequence 8, Appli  |
| 42 | 71   | 9.9  | 449  | 16 | US-10-437-963-158433 | Sequence 158433, A |
| 43 | 71   | 9.9  | 645  | 16 | US-10-437-963-125813 | Sequence 125813, A |
| 44 | 71   | 9.9  | 707  | 16 | US-10-437-963-142238 | Sequence 142238, A |
| 45 | 70.5 | 9.8  | 360  | 14 | US-10-195-144-9      | Sequence 9, Appli  |

#### ALIGNMENTS

#### RESULT 1

US-09-943-002-14  
; Sequence 14, Application US/09943002  
; Patent No. US20020045734A1  
; GENERAL INFORMATION:  
; APPLICANT: Duncan, Roy  
; TITLE OF INVENTION: NOVEL REOVIRUS-DERIVED PROTEINS AND USES THEREFOR  
; FILE REFERENCE: 78973-1C  
; CURRENT APPLICATION NUMBER: US/09/943,002  
; CURRENT FILING DATE: 2001-08-31  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 14  
; LENGTH: 140  
; TYPE: PRT  
; ORGANISM: baboon reovirus  
US-09-943-002-14

Query Match 100.0%; Score 720; DB 9; Length 140;  
Best Local Similarity 100.0%; Pred. No. 3.7e-72;  
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|    |     |                                                            |     |
|----|-----|------------------------------------------------------------|-----|
| QY | 1   | MGQRHSIVQPPAPPNAFVEIVSSSTGIIIAVGIPAFIFSLYKLLQWYNRKSKNRKKE  | 60  |
| DB | 1   | MGQRHSIVQPPAPPNAFVEIVSSSTGIIIAVGIPAFIFSLYKLLQWYNRKSKNRKKE  | 60  |
| QY | 61  | QIRQIEIGLLSYGAGVASLPLLVIAHNPQSVISATPIYKGPCTGVPSNRLLOITSGTA | 120 |
| DB | 61  | QIRQIEIGLLSYGAGVASLPLLVIAHNPQSVISATPIYKGPCTGVPSNRLLOITSGTA | 120 |
| QY | 121 | BENTRILNHDGRNPDGSINV                                       | 140 |
| DB | 121 | BENTRILNHDGRNPDGSINV                                       | 140 |

#### RESULT 2







QY 105 T--GVNSR-----LLQITSGTAREN 123  
DB 389 EPGVSSRSWFYEBLQATNGFSAQN 416

## RESULT 11

US-10-425-115-290480  
; Sequence 290480, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 290480  
; LENGTH: 539  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_28005C.1.pep  
US-10-425-115-290480

Query Match 11.0%; Score 79.5; DB 17; Length 539;  
Best Local Similarity 35.6%; Pred. No. 8.9; Indels 1; Gaps 1;  
Matches 21; Conservative 10; Mismatches 27

QY 81 PLLNVAHPGVSATPIYKGPCTGVNSRLLOITSGTAENTRLNHGDR-NPDGSI 138  
DB 8 PLLVNNHSPGPVATPGEHGAAYDESSTLVDAASGTDESHRALADADVKNDDGDV 66

## RESULT 12

US-10-425-114-43526  
; Sequence 43526, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 43526  
; LENGTH: 544  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700444335\_FLI.pep  
US-10-425-114-43526

Query Match 11.0%; Score 79.5; DB 15; Length 544;  
Best Local Similarity 35.6%; Pred. No. 9; Indels 1; Gaps 1;  
Matches 21; Conservative 10; Mismatches 27

QY 81 PLLNVAHPGVSATPIYKGPCTGVNSRLLOITSGTAENTRLNHGDR-NPDGSI 138  
DB 13 PLLVNNHSPGPVATPGEHGAAYDESSTLVDAASGTDESHRALADADVKNDDGDV 71

## RESULT 13

US-10-437-963-183058

; Sequence 183058, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 183058  
; LENGTH: 419  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_80186C.1.pep  
US-10-437-963-183058

Query Match 10.9%; Score 78.5; DB 16; Length 419;  
Best Local Similarity 23.0%; Pred. No. 8.3; Indels 41; Gaps 7;  
Matches 32; Conservative 22; Mismatches 44

QY 13 PPNAPVEIVSSGTGIIIAVGIFAFIPFLYKLLWYNRKSNKKGKQIREQI--ELGL 70  
DB 239 PPPIAHLHAKA-----IMELWYNNGYKYOMESRSNOGIMEEDSV 281

QY 71 LSVGAGVASLPILNVAHPGVSATPIYKGPCTGVNSRLLOI-----TSGTAEN 123  
DB 282 ISNGREBEV-----VWANGSGGVGGT-----ARSSGGGGKVVQVLQRNPFGEVQILEQN 333

QY 124 TRIL-----NHDGRNPDG 136  
DB 334 -RVLIQELSQNHARDADG 351

## RESULT 14

US-10-369-493-10884  
; Sequence 10884, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 10884  
; LENGTH: 325  
; TYPE: PRT  
; ORGANISM: Sphingomonas aromaticivorans  
US-10-369-493-10884

Query Match 10.8%; Score 78; DB 14; Length 325;  
Best Local Similarity 36.7%; Pred. No. 6.7; Indels 0; Gaps 0;  
Matches 18; Conservative 7; Mismatches 24

QY 90 PGVSATPIYKGPCTGVNSRLLOITSGTAENTRLNHGDRNPDGSI 138  
DB 92 PGAAIHAQPSYRTHGEMIPRSRPVALTBLTGTEATILANPNPDGRI 140



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM.protein - protein search, using sw model

Run on: November 15, 2004, 14:46:58 ; Search time 38 Seconds  
(without alignments)  
354.483 Million cell updates

**Title:** US-09-943-002-14

Perfect score: 720  
Sequence: 1 MGQRHSIVQPPAPPNFAVE.....EENTRIIHDGRNPDGSINV 140

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

Database : PIR 79:★

```

1:  p1r1:
2:  p1r2:
3:  p1r3:
4:  p1r4:

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID     | Description         |
|------------|-------|-------------|--------|----|--------|---------------------|
| 1          | 84    | 11.7        | 431    | 2  | C89979 | adenylosuccinate 1  |
| 2          | 83    | 11.5        | 167    | 2  | D70401 | phosphoribosylamin  |
| 3          | 78    | 10.8        | 255    | 2  | A75507 | conserved hypothet  |
| 4          | 78    | 10.8        | 391    | 2  | AH1328 | cell-division init  |
| 5          | 78    | 10.8        | 392    | 2  | AI1699 | cell-division init  |
| 6          | 77.5  | 10.8        | 339    | 2  | A81351 | signal transduction |
| 7          | 77.5  | 10.8        | 654    | 2  | C87791 | protein B0207.12 [  |
| 8          | 76    | 10.6        | 298    | 2  | T36900 | probable integral   |
| 9          | 76    | 10.6        | 1477   | 2  | T19534 | protein-tyrosine k  |
| 10         | 75.5  | 10.5        | 427    | 2  | S19338 | hypothetical prote  |
| 11         | 75.5  | 10.5        | 1220   | 2  | S48387 | SLN1 protein - yea  |
| 12         | 75    | 10.4        | 399    | 1  | S76745 | hypothetical prote  |
| 13         | 74.5  | 10.3        | 3005   | 2  | D90031 | hypothetical prote  |
| 14         | 74.5  | 10.3        | 359    | 2  | T36781 | probable gluconate  |
| 15         | 74    | 10.3        | 433    | 2  | S05654 | ND3 intron protein  |
| 16         | 74    | 10.3        | 468    | 2  | T33857 | hypothetical prote  |
| 17         | 74    | 10.3        | 599    | 2  | T15774 | hypothetical prote  |
| 18         | 73.5  | 10.2        | 256    | 2  | T45049 | hypothetical prote  |
| 19         | 73    | 10.1        | 443    | 2  | S11396 | gamma-aminobutyric  |
| 20         | 73    | 10.1        | 3036   | 2  | T18995 | hypothetical prote  |
| 21         | 72.5  | 10.1        | 120    | 2  | S07727 | NADH2 dehydrogenas  |
| 22         | 72.5  | 10.1        | 133    | 2  | T30474 | hypothetical prote  |
| 23         | 72.5  | 10.1        | 338    | 2  | T23260 | hypothetical prote  |
| 24         | 72.5  | 10.1        | 1677   | 2  | T46095 | hypothetical prote  |
| 25         | 72    | 10.0        | 453    | 2  | S11087 | gamma-aminobutyric  |
| 26         | 72    | 10.0        | 631    | 2  | D63750 | phosphotransferase  |
| 27         | 72    | 10.0        | 714    | 2  | AF3479 | ABC transporter AT  |
| 28         | 72    | 10.0        | 910    | 2  | A53137 | tyrosine kinase re  |
| 29         | 71.5  | 9.9         | 210    | 2  | G97235 | hypothetical prote  |

|    |      |     |      |   |        |                                   |
|----|------|-----|------|---|--------|-----------------------------------|
| 30 | 71.5 | 9.9 | 385  | 2 | AH3477 | tryptophan-tRNA ligase            |
| 31 | 70.5 | 9.8 | 431  | 2 | A72294 | adenylosuccinate lyase            |
| 32 | 70.5 | 9.8 | 445  | 2 | B83931 | hypothetical protein              |
| 33 | 70.5 | 9.8 | 1588 | 2 | T38660 | probable transcription factor     |
| 34 | 70   | 9.7 | 323  | 2 | S61927 | lipase A precursor                |
| 35 | 70   | 9.7 | 536  | 2 | A55544 | flagellar basal body              |
| 36 | 70   | 9.7 | 708  | 2 | D96711 | hypothetical protein              |
| 37 | 69.5 | 9.7 | 358  | 2 | F71110 | hypothetical protein              |
| 38 | 69.5 | 9.7 | 557  | 2 | S60392 | phosphoribosylamin                |
| 39 | 69.5 | 9.7 | 881  | 2 | T33810 | hypothetical protein              |
| 40 | 69.5 | 9.7 | 1119 | 2 | T32074 | hypothetical protein              |
| 41 | 69   | 9.6 | 432  | 2 | AE1934 | periplasmic sugar-binding protein |
| 42 | 69   | 9.6 | 500  | 1 | JE0259 | cytochrome P450 2D                |
| 43 | 69   | 9.6 | 502  | 2 | AD0691 | conserved hypothetical protein    |
| 44 | 69   | 9.6 | 589  | 2 | E95045 | PTS system, mannitol              |
| 45 | 69   | 9.6 | 589  | 2 | D97916 | phosphotransferase                |

## ALIGNMENTS

## RESULT 1

adenylosuccinate lyase [imported] - *Staphylococcus aureus* (strain N315)  
 C:Species: *Staphylococcus aureus*  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
 C:Accession: AB09799  
 R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.;  
 ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; S  
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
 Lancet 357, 1225-1240, 2001  
 A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aure*  
 A:Reference number: AB9758; UID:21311952; PMID:11418146

A;Accession: C89979  
A;Status: preliminary

A;Molecule type: DNA  
A;Residues: 1-431 <KUR>

A;Residues: 1-431 (R0K)  
A;Cross-references: UNIPROT:Q99SX9  
A;Experimental source: strain N315

A; Experimental source: strain N313  
C: Genetics:

A:Gene: purB

C;Superfamily: fumarate hydratase

Query Match: 11.7% Score 84: DB 2: Length 431:

|                       |       |                |    |
|-----------------------|-------|----------------|----|
| Query Match           | 11.7% | Score 64;      | DB |
| Best Local Similarity | 26.5% | Pred. No. 1.8; |    |

| Desc    | Local Similarity | 20:50, Freq. NO. | 1:0,           |
|---------|------------------|------------------|----------------|
| Matches | 26:              | Conservative     | 15: Mismatches |
|         | 22:              | Indels           | 35: Gaps       |

33 GIPAFIFSFYKLLWYNRKSQKRRKEOIREOIELGLLSYGAGVASLPLLNVIAHNPGS 92

[illegible]

93 VISATPIYGCPCWCIVNSPLIOTISCTABENTRI.NHD 130

|     |                                        |
|-----|----------------------------------------|
| 93  | VISATPIYKGFCTGVFNSRLLQIISGTABENIKILNHD |
| :   | : : : : :                              |
| 184 | TGVSIVCTC-----TCTA DYSTOTI QDD         |
| 217 | -----                                  |

## RESULT 2

phosphoribosylaminoimidazole carboxylase (EC 4.1.1.21) catalytic chain - Aquifex aeolicus  
C;Species: Aquifex aeolicus  
C;Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jul-2004  
C;Accession: D70401

C/Accession: D/0401  
R/Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V

Nature 392, 353-358, 1998  
A. Title. The complete genome of the hyperthermophilic bacterium *Amiifex azo* [Jus. V.

A;Title: The complete genome of the hyperthermophilic  
A;Reference number: A70300: MUID:98196666: PMID:9537320

A:Reference number: A/0300; MUID:98196666; PMID:9337320  
A:Accession: D70401

A;Accession: D/0401  
A;Status: preliminary

A: Molecule type: DNA

A: Molecule type: DNA  
A: Residues: 1-167 <AOF>

A:Accession: 1-107 (SAP)  
A:Cross-references: UNIPROT:O67239; GB:AE000727; NID:g2983623; PIDN:AAC07201.1; PID:g291  
A:Cross-references: UNIPROT:O67239; GB:AE000727; NID:g2983623; PIDN:AAC07201.1; PID:g291  
A:Experimental source: strain VF5



C:Genetics:  
A:Gene: purE  
C:Superfamily: phosphoribosylaminoimidazole carboxylase catalytic chain; phosphoribosyl  
C:Keywords: carbon-carbon lyase; carboxy-lyase; purine nucleotide biosynthesis  
P:3-136/Domain: phosphoribosylaminoimidazole carboxylase catalytic chain homology <PC>  
Query Match 11.5%; Score 83; DB 2; Length 167;  
Best Local Similarity 33.8%; Pred. No. 0.77;  
Matches 23; Conservative 15; Mismatches 28; Indels 2; Gaps 1;  
QY 5 HSTVQPPAPPNAPFVIVSSGIIITAVGIFAFIFSFYKLLQWYNRKSNK--KKKEQI 62  
DB 100 YSLVQNPAGIVATVAIGNATNAGLLAVRLISIKYPEYAKKLDXYTEKLKRVKAGNSEL 159  
QY 63 REQIELGL 70  
DB 160 QKEVENGI 167  
RESULT 3  
A:Map position: 1  
C:Superfamily: hypothetical protein - Deinococcus radiodurans (strain R1)  
conserved hypothetical protein - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: A75507  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, C.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Pshihi, H.; Jones, L.M.; Karst, U.; Science 294, 849-852, 2001  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896; PMID:10567266  
A:Accession: A75507  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-255 <WHI>  
A:Cross-references: UNIPROT:Q9RWX1; GB:AE001912; GB:AE000513; NID:96458228; PIDN:AAF1011  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR0544  
A:Map position: 1  
C:Superfamily: hypothetical protein HI0902  
Query Match 10.8%; Score 78; DB 2; Length 255;  
Best Local Similarity 27.3%; Pred. No. 3.9;  
Matches 30; Conservative 14; Mismatches 42; Indels 24; Gaps 3;  
QY 22 VSSSTGIIIVGIFAFIPFSL-----YKLLQWYNRKSNKKEQIREQIEL----- 68  
DB 94 VGSFGLVLPARAVATVFALLFYSAYNLL-----RGLKRVVEREPEKLVPPAMTPA 146  
QY 69 ---GLLSYGAGVASLPLLVIAHNPVSGVISATPIYKGPCTGVPSRLAQ 114  
DB 147 GWSGLLGIGGTQVQPVVLNLAQGMPIRQAIATSTFIMGLTAVGNALVTQ 196  
RESULT 4  
AH1328  
cell-division initiation protein FtsZ homolog ftsZ [imported] - Listeria monocytogenes  
C:Species: Listeria monocytogenes  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: AH1328  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, C.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Pshihi, H.; Jones, L.M.; Karst, U.; Science 294, 849-852, 2001  
A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AH1328  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-391 <GLA>

A:Cross-references: UNIPROT:Q8YSM5; GB:NC\_003210; PIDN:CAD00110.1; PID:g16411502; GSPDB  
A:Experimental source: strain EGD-e  
C:Genetics:  
A:Gene: ftsZ  
C:Superfamily: cell division protein ftsZ  
Query Match 10.8%; Score 78; DB 2; Length 391;  
Best Local Similarity 28.3%; Pred. No. 6.3;  
Matches 30; Conservative 14; Mismatches 24; Indels 38; Gaps 5;  
QY 56 KKRKEQIREQIELGL-----LSYGAGVASLPLLVIAHNPVSGVISATPIY 100  
DB 79 KKAABESREQIEALKGDVMTAGMGGGTGTGAAPVIAQIAKEMGALTGVGVTTPPGF 138  
QY 101 KGP-----CTG-----VPNSRLLOITSGTAENTRIL 127  
DB 139 EGPKRKTKALGTGTEAMKEAVDTLIVIPNDRLLQI-----VDKNTPLM 180  
RESULT 5  
AH1699  
cell-division initiation protein FtsZ homolog ftsZ [imported] - Listeria innocua (strain  
C:Species: Listeria innocua  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: AH1699  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, C.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Pshihi, H.; Jones, L.M.; Karst, U.; Science 294, 849-852, 2001  
A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AH1699  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-392 <GLA>  
A:Cross-references: UNIPROT:Q929Y5; GB:AL592022; PIDN:CAC97368.1; PID:g16414652; GSPDB:  
A:Experimental source: strain Clip11262  
C:Genetics:  
A:Gene: ftsZ  
C:Superfamily: cell division protein ftsZ  
Query Match 10.8%; Score 78; DB 2; Length 392;  
Best Local Similarity 28.3%; Pred. No. 6.4;  
Matches 30; Conservative 14; Mismatches 24; Indels 38; Gaps 5;  
QY 56 KKRKEQIREQIELGL-----LSYGAGVASLPLLVIAHNPVSGVISATPIY 100  
DB 79 KKAABESREQIEALKGDVMTAGMGGGTGTGAAPVIAQIAKEMGALTGVGVTTPPGF 138  
QY 101 KGP-----CTG-----VPNSRLLOITSGTAENTRIL 127  
DB 139 EGPKRKTKALGTGTEAMKEAVDTLIVIPNDRLLQI-----VDKNTPLM 180  
RESULT 6  
AH1351  
signal transduction histidine kinase Cj0793 [imported] - Campylobacter jejuni (strain N  
C:Species: Campylobacter jejuni  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C:Accession: AB1351  
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Baeham, D.; Chilli, C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barre, Nature 403, 663-668, 2000  
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hy  
A:Reference number: AB1250; MUID:20150912; PMID:10688204  
A:Accession: AB1351  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-339 <PAR>  
A:Cross-references: UNIPROT:Q9PPC6; GB:AL139076; GB:AL111168; NID:g6968128; PIDN:CAB730;  
A:Experimental source: serotype O2, strain NCTC 11168



| Matches | 31, | Conservative                                                  | 18, | Mismatches | 74; | Indels | 9; | Gaps | 3, |
|---------|-----|---------------------------------------------------------------|-----|------------|-----|--------|----|------|----|
| Qy      | 4   | RHSIVQPPAPPNPAFVIVSSSTGIIAVGFATFISFLYKLLQWNRKSNKKRKEQIR       | 63  |            |     |        |    |      |    |
| Db      | 189 | RCLMVRQDNPPPEAAKFSKATREEDNGFEVFIINLPYB-MNWQSLKDMFKECGHVLR     | 247 |            |     |        |    |      |    |
| Qy      | 64  | EQTELGLLSYGAGVASLPLNLVIAHNFGSVISATPIYKGPCTGVPNGRLLIQITSGTAAEN | 123 |            |     |        |    |      |    |
| Db      | 248 | ADVELDPNGFSRGRGS-----VIYPTDEMIRAIDTFNGM---EVEGRVLEVRGGRFNKR   | 299 |            |     |        |    |      |    |
| Qy-     | 124 | TRILNHDGRNPD                                                  | 135 |            |     |        |    |      |    |
| Db      | 300 | KNNDRYNQREED                                                  | 311 |            |     |        |    |      |    |

RESULT 11  
 S48387  
 SLN1 protein - yeast (*Saccharomyces cerevisiae*)  
 N;Alternate names: protein Y1L147C  
 C;Species: *Saccharomyces cerevisiae*  
 C;Date: 02-Dec-1994 #sequence\_revision 02-Dec-1994 #text\_change 09-Jul-2004  
 C;Accession: S48387; S48888  
 R;Churcher, C.  
 submitted to the EMBL Data Library, September 1994  
 A;Reference number: S48310  
 A;Accession: S48387  
 A;Molecule type: DNA  
 A;Residues: 1-1220 <CHU>  
 A;Cross-references: UNIPROT:P39928; GB:Z47047; EMBL:Z38059; NID:G603937; PID:G763139; MI:G763139; PI:G763139  
 R;Ota, I.M.; Varshavsky, A.  
 Science 262, 566-569, 1993  
 A;Title: A yeast protein similar to bacterial two-component regulators.  
 A;Reference number: S48888; MUID:94024030; PMID:8211183  
 A;Accession: S48888  
 A;Status: nucleic acid sequence not shown  
 A;Molecule type: DNA  
 A;Residues: 1-1220 <OTA>  
 A;Cross-references: EMBL:U01835; NID:G414473; PIDN:ACC48912.1; PID:G414474  
 C;Genetics:  
 A;Gene: SGD:SLN1; YPD2  
 A;Cross-references: SGD:S0001409; MIPS:Y1L147C  
 A;Map position: 9L  
 F;24-40/Domain: transmembrane #status predicted <TML>  
 F;334-350/Domain: transmembrane #status predicted <TM2>  
 F;1096-1206/Domain: response regulator homology <RRH>  
 F;1144/Binding site: phosphate (Asp) (covalent) #status predicted

```

Query Match      10.5%; Score 75.5; DB 2; Length 1220;
Best Local Similarity 23.3%; Pred. No. 41;
Matches 31; Conservative 22; Mismatches 53; Indels 27; Gaps 6

QY      1  MGRHSIVQPPAPPNAPVEIVSSGTGIIITAVGIFAFIPFLYKLLQWYNRKSKKKRE 60
      .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
DB      317  VSQPSVFLSPA-----TKLAKIITGVIAIGVFVILLTL--PLAHW-----AVOP 360

QY      61  QIREQIBELGLLSYGAGVASLPLLNVIAHNPQGSVSATPIYKPGCTG--VPNSRLLQITSG 118
      .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
DB      361  IVRLOKATLITEGRLR-----PSTPTTIRASSFKRSGSGFANPSS--LLQFNPTA 411

QY      119  TABENTRIINHDG 131

DB      412  EAGSTTSVSGHG 424

```

RESULT 12  
S76745  
hypothetical protein - *Synechocystis* sp. (strain PCC 6803)  
C.Species: *Synechocystis* sp.  
A.Variety: PCC 6803  
C.Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C.Accession: S76745  
R.Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.

o, K.; Okumura, S.; Shimpō, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
DNA Res. 3, 109-136, 1996  
A/Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*  
8.  
A/Reference number: S74322; MUID:97061201; PMID:8905231  
A/Accession: S76745  
A/Status: nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-399 <KAN>  
A/Cross-references: UNIPROT:P74550; EMBL:D90916; GB:AB001339; NID:g1653715; PIDN:BAAL86  
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C/Genetics:  
A/Start codon: GTG  
C/Superfamily: Chromate resistance protein A

```

Query Match      10.4%; Score 75; DB 1; Length 399;
Best Local Similarity 23.8%; Pred. No. 13;
Matches 24; Conservative 23; Mismatches 34; Indels 20; Gaps 4;

QY 12 APPNPAFVIVS--STGIIIAVG-----IPAFISFLYKLOWYRKSKNKKRKE 60
   |||  ::  |||  ::  |||  ::  |||  ::  |||  ::  |||  ::
Db 214 APAPALPVTIINPWSVFGIFLKGILYGGVYLFAPLQADLVERLEWLT-----SQ 265

QY 61 QIREQIELGLLYSGACVASLPILN-VIAHNPGSVISATPIY 100
   ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
Db 266 QLLDIAIGCIGTPGLPTATTGTYILLAGNPGAAIATGVIF 306

```

RESULT 13  
D90031  
hypothetical protein S2112 [imported] - Staphylococcus aureus (strain N315)  
C:Species: Staphylococcus aureus  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C:Accession: D90031  
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaico, C.; Sekimizu, K.; Ma, C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayaishi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001  
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
A:Reference number: A89758; PMID:21311952; PMID:11418146  
A:Accession: D90031  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-305 <Kur>  
A:Cross-references: UNIPROT:Q99RV2; GB:BA000018; PID:g13702121; PIDN:BAB43413.1; GSPDB: A15  
A:Experimental source: strain N315

[illegible]

```

RESULT 14
T36781
      probable gluconate permease - Streptomyces coelicolor (fragment)
C-Species: Streptomyces coelicolor
C.Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change
C.Accession: T36781
R:J.Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell,
submitted to the EMBL Data Library, July 1999
A:Reference number: 221614
A:Accession: T36781

```

Search completed: November 15, 2004, 14:58:22  
Job time : 45 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 15, 2004, 14:32:21 ; Search time 195 Seconds  
(without alignments)

413.089 Million cell updates/sec

Title: US-09-943-002-14

Perfect score: 720

Sequence: 1 MGRHSIVQPPPPNFAVE.....EENTRIINHGRNPDGSINV 140

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt\_02:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID           | Description        |
|------------|-------|-------------|--------|--------------|--------------------|
| 1          | 720   | 100.0       | 140    | 2 Q918V6     | Q918V6 baboon reov |
| 2          | 85.5  | 11.9        | 806    | 2 Q7QPH4     | Q7QPH4 giardia lam |
| 3          | 84    | 11.7        | 431    | 2 Q6G825     | Q6G825 staphylococ |
| 4          | 84    | 11.7        | 431    | 2 Q6GFE9     | Q6GFE9 staphylococ |
| 5          | 84    | 11.7        | 431    | 2 Q998X9     | Q998X9 staphylococ |
| 6          | 84    | 11.7        | 431    | 2 Q7A0G9     | Q7A0G9 staphylococ |
| 7          | 84    | 11.7        | 431    | 2 Q7A4Q3     | Q7A4Q3 staphylococ |
| 8          | 84    | 11.7        | 574    | 2 Q8DQC1     | Q8DQC1 synchococ   |
| 9          | 83    | 11.5        | 167    | 1 PURE AQUAE | Q67239 aquifex ae  |
| 10         | 82.5  | 11.5        | 536    | 2 Q22040     | Q22040 caenorhabdi |
| 11         | 82    | 11.4        | 734    | 2 Q7Q696     | Q7Q696 anopheles g |
| 12         | 80.5  | 11.2        | 431    | 2 Q8CRTE     | Q8CRTE staphylococ |
| 13         | 79.5  | 11.0        | 443    | 2 Q6F2A9     | Q6F2A9 mesoplasma  |
| 14         | 79.5  | 11.0        | 1491   | 2 Q8C115     | Q8C115 m mus muscu |
| 15         | 79    | 11.0        | 633    | 2 Q8LF72     | Q8LF72 nicotiana t |
| 16         | 79    | 11.0        | 664    | 2 Q899Q9     | Q899Q9 clostridium |
| 17         | 78    | 10.8        | 255    | 2 Q9RWX1     | Q9RWX1 deinococcus |
| 18         | 78    | 10.8        | 391    | 2 Q8Y5M5     | Q8Y5M5 listeria mo |
| 19         | 78    | 10.8        | 391    | 2 Q71XY1     | Q71XY1 listeria mo |
| 20         | 78    | 10.8        | 391    | 2 AAT04834   | AAT04834 listeria  |
| 21         | 78    | 10.8        | 392    | 2 Q929Y5     | Q929Y5 listeria in |
| 22         | 77.5  | 10.8        | 339    | 2 Q9PPC6     | Q9PPC6 campylobact |
| 23         | 77.5  | 10.8        | 430    | 2 Q17369     | Q17369 caenorhabdi |
| 24         | 77.5  | 10.8        | 430    | 2 Q95Q97     | Q95Q97 caenorhabdi |
| 25         | 77    | 10.7        | 172    | 2 Q83314     | Q83314 enterococcu |
| 26         | 76.5  | 10.6        | 681    | 2 Q9FEW5     | Q9FEW5 arabidopsis |
| 27         | 76    | 10.6        | 187    | 2 Q8RF10     | Q8RF10 fusobacteri |
| 28         | 76    | 10.6        | 280    | 2 Q6YT74     | Q6YT74 oryza sativ |
| 29         | 76    | 10.6        | 280    | 2 BAC84718   | BAC84718 oryza sat |
| 30         | 76    | 10.6        | 298    | 2 Q9X9Y6     | Q9X9Y6 streptomyce |
| 31         | 76    | 10.6        | 308    | 2 Q8TQ34     | Q8TQ34 methanosarc |

|    |      |      |      |              |                     |
|----|------|------|------|--------------|---------------------|
| 32 | 76   | 10.6 | 765  | 2 Q7M857     | Q7M857 wolinnella s |
| 33 | 76   | 10.6 | 1477 | 1 HTK7 HYDAT | Q25197 hydra atten  |
| 34 | 75.5 | 10.5 | 427  | 1 GBP2 YEAST | P25555 saccharomyc  |
| 35 | 75.5 | 10.5 | 1220 | 1 SLN1 YEAST | P39928 saccharomyc  |
| 36 | 75   | 10.4 | 399  | 2 P74550     | P74550 synchococyt  |
| 37 | 75   | 10.4 | 481  | 2 Q7RG91     | Q7RG91 plasmodium   |
| 38 | 75   | 10.4 | 508  | 2 Q722R7     | Q722R7 bacillus ce  |
| 39 | 75   | 10.4 | 508  | 2 AAS43502   | AAS43502 bacillus   |
| 40 | 75   | 10.4 | 626  | 1 LPXB CHLCV | Q82123 chlamydomon  |
| 41 | 75   | 10.4 | 848  | 2 Q8P247     | Q8P247 methanosarc  |
| 42 | 75   | 10.4 | 892  | 2 Q9Y438     | Q9Y438 homo sapien  |
| 43 | 74.5 | 10.3 | 305  | 2 Q8G629     | Q8G629 staphylococ  |
| 44 | 74.5 | 10.3 | 305  | 2 Q8NV71     | Q8NV71 staphylococ  |
| 45 | 74.5 | 10.3 | 305  | 2 Q99RV2     | Q99RV2 staphylococ  |

ALIGNMENTS

RESULT 1

Q918V6 PRELIMINARY; PRT; 140 AA.  
AC Q918V6;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE Membrane fusion protein p15.  
OS Baboon reovirus.  
OC Viruses; dsRNA viruses; Reoviridae; Orthoreovirus.  
OX NCBI\_TaxID=75888;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Dawe S., Duncan R.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Shou J., Chen Z., Duncan R.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF406787; AAL01373.1; -  
SQ SEQUENCE 140 AA; 15222 MW; B02BDF37B6CE6075 CRC64;

Query Match 100.0%; Score 720; DB 2; Length 140;  
Best Local Similarity 100.0%; Pred. No. 1.7e-63;  
Matches 140; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

|    |     |                                                             |     |
|----|-----|-------------------------------------------------------------|-----|
| Qy | 1   | MGQRHSIVQPPPPNFAVEIVSSSTGIIIVGIFAFISFLYKLLQWYRKSNKKRKE      | 60  |
| Db | 1   | MGQRHSIVQPPPPNFAVEIVSSSTGIIIVGIFAFISFLYKLLQWYRKSNKKRKE      | 60  |
| Qy | 61  | QIREQIEIGLLSVGAGVASLPLLNVIHNPVGSVISATPIYKGPCTGVPNSRLQITSCTA | 120 |
| Db | 61  | QIREQIEIGLLSVGAGVASLPLLNVIHNPVGSVISATPIYKGPCTGVPNSRLQITSCTA | 120 |
| Qy | 121 | EENTRIINHGRNPDGSINV 140                                     |     |
| Db | 121 | EENTRIINHGRNPDGSINV 140                                     |     |

RESULT 2

Q7QPH4 PRELIMINARY; PRT; 806 AA.  
AC Q7QPH4;  
DT 01-MAR-2004 (TREMBlrel. 26, Created)  
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)  
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
DE GLP 41 11524 13944.  
OS Giardia lamblia ATCC 50803.  
OC Eukaryota; Diplomonadida; Hexamitidae; Giardinae; Giardia.  
OX NCBI\_TaxID=184922;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WB C6;  
RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,

Olsen G.J., Sogin M.L.;  
 "Draft sequence of the Giardia lamblia genome.";  
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AAC801000.169; EAA36911.1; -;  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR001965; Znf\_PHD.  
 DR Pfam; PF00628; PHD; 1.  
 SU SEQUENCE 806 AA; 90537 MW; CD93B0653EF38DB9 CRC64;

Query Match 11.9%; Score 85.5; DB 2; Length 806;  
 Best Local Similarity 24.5%; Pred. No. 18;  
 Matches 35; Conservative 21; Mismatches 44; Indels 43; Gaps 6;

QY 14 PPNAFVEIVSSSTGIIIAVGFAIFSFYKLLWYNRKSKKKKEQIRQIEGLLSY 73  
 Db 387 PPNFPMKVLRD-----PIRLREIQOMKALQWYGLLKQVTSPEKSSPEPKDQ 440  
 QY 74 GAGVASLPDLLNVAHNPGSVISATPIYKPGCTGV-----PNSRLQ-----ITSGT 119  
 Db 441 GHGVLAFLPVLTL-----SSIIIGSGI---RTGVVISINDITSANSRLRLEAGTNGDMTPGS 492  
 QY 120 AE-----ENTRIL 127  
 Db 493 ASAANGPHASMRQTSNKTKVI 515

RESULT 3  
 Q6G825 PRELIMINARY; PRT; 431 AA.  
 ID Q6G825  
 AC Q6G825;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Adenylosuccinate lyase (EC 4.3.2.2).  
 GN ORFNames=SAS1831;  
 OS Staphylococcus aureus subsp. aureus MSSA476.  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 NCBI\_TaxID=282459;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MSSA476;  
 RA Holden M.T.G., Feil E.J., Lindsay J.A., Peacock S.J., Day N.P.J.,  
 RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,  
 RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,  
 RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,  
 RA Fellwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,  
 RA James K.D., Lennard N., Line A., Mayes R., Moule S., Mungall K.,  
 RA Ormond D., Quail M.A., Rabinowitsch E., Rutherford K., Sanders M.,  
 RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,  
 RA Spratt B.G., Parkhill J.;  
 RA "Complete genomes of two clinical Staphylococcus aureus strains:  
 RT evidence for the rapid evolution of virulence and drug resistance.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).  
 DR EMBL; BX571856; CAG40985.1; -;  
 DR InterPro; IPR003031; D\_crySTALLIN.  
 DR InterPro; IPR000362; Fumarate lyase.  
 DR InterPro; IPR008948; L-Aspartase-like.  
 DR InterPro; IPR004769; Pur lyase.  
 DR Pfam; PF00206; Lyase 1; I.  
 DR PRINTS; PR00145; DCRYSTALLIN.  
 DR PRINTS; PR00149; FUMARATELYASE.  
 DR TIGRPFAMS; TIGR00928; purB; 1.  
 DR PROSITE; PS00163; FUMARATE\_LYASES; 1.  
 KW Lyase.  
 SQ SEQUENCE 431 AA; 49603 MW; 493F79CBE814B9E5 CRC64;

Query Match 11.7%; Score 84; DB 2; Length 431;  
 Best Local Similarity 26.5%; Pred. No. 12;  
 Matches 26; Conservative 15; Mismatches 35; Indels 22; Gaps 2;

QY 33 GIFAFIFSYKLLWYNRKSKKKKEQIRQIEGLLSYGAGVASLPDLLNVAHNPGS 92  
 Db 142 GVHAEPITFGVQKALWYTEMQRLQRFKQVREIEVGMKSGAVG-----TFANIPPE 193

RESULT 5  
 Q99SX9 PRELIMINARY; PRT; 431 AA.  
 ID Q99SX9  
 AC Q99SX9;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Adenylosuccinate lyase.  
 GN Names=purb; OrderedLocusNames=SAV1908;  
 OS Staphylococcus aureus (strain Mu50 / ATCC 700699).  
 NCBI\_TaxID=282458;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MSSA252;  
 RA Holden M.T.G., Feil E.J., Lindsay J.A., Peacock S.J., Day N.P.J.,  
 RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,  
 RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,  
 RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,  
 RA Fellwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,  
 RA James K.D., Lennard N., Line A., Mayes R., Moule S., Mungall K.,  
 RA Ormond D., Quail M.A., Rabinowitsch E., Rutherford K., Sanders M.,  
 RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,  
 RA Spratt B.G., Parkhill J.;  
 RA "Complete genomes of two clinical Staphylococcus aureus strains:  
 RT evidence for the rapid evolution of virulence and drug resistance.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).  
 DR EMBL; BX571857; CAG43636.1; -;  
 DR InterPro; IPR003031; D\_crySTALLIN.  
 DR InterPro; IPR000362; Fumarate lyase.  
 DR InterPro; IPR008948; L-Aspartase-like.  
 DR InterPro; IPR004769; Pur lyase.  
 DR Pfam; PF00206; Lyase 1; I.  
 DR PRINTS; PR00145; DCRYSTALLIN.  
 DR PRINTS; PR00149; FUMARATELYASE.  
 DR TIGRPFAMS; TIGR00928; purB; 1.  
 DR PROSITE; PS00163; FUMARATE\_LYASES; 1.  
 KW Lyase.  
 SQ SEQUENCE 431 AA; 49603 MW; 493F79CBE814B9E5 CRC64;

Query Match 11.7%; Score 84; DB 2; Length 431;  
 Best Local Similarity 26.5%; Pred. No. 12;  
 Matches 26; Conservative 15; Mismatches 35; Indels 22; Gaps 2;

```

OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NU50 / ATCC 700699;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita N., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RA "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AB003363; BAB58070.1; --
DR FIC; CB9979; CB9979.
DR GO; GO:0004018; F:adenylosuccinate lyase activity; IEA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0009152; P:purine ribonucleotide biosynthesis; IEA.
DR InterPro; IPR003031; D:crystallin.
DR InterPro; IPR000362; Fumarate_lyase.
DR InterPro; IPR008948; L-Aspartase-like.
DR Pfam; PF00206; Lyase 1; 1.
DR PRINTS; PR00145; DCRYSTALLIN.
DR PRINTS; PR00149; FUMATELYASE.
DR TIGRPFAMs; TIGR00928; purB; 1.
DR Complete proteome; IEA.
KW Complete proteome; Lyase.
SQ SEQUENCE 431 AA; 49603 MW; 493F79CBE814B9E5 CRC64;

Query Match 11.7%; Score 84; DB 2; Length 431;
Best Local Similarity 26.5%; Pred. No. 12;
Matches 26; Conservative 15; Mismatches 35; Indels 22; Gaps 2;

QY 33 GIFAIFSFYKLLQWNRKSKNKKRKEQIRBQIEGLLSYGAGVASLPLNVIHNPQS 92
DB 142 GVHAEPPTFGVKWALWYTEMQNLQRFQVREIEVGKMSGAVG-----TFANIPPE 193

QY 93 VISATPIYKGPCTGVPSNRLQITSGTAENTRILNHD 130
DB 194 IESYVCKHLG-----IGTAPVSTQTLQRD 217

RESULT 6
Q7A0G9 PRELIMINARY; PRT; 431 AA.
AC Q7A0G9
DT 05-JUL-2004 (TREMELrel. 27, Created)
DT 05-JUL-2004 (TREMELrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMELrel. 27, Last annotation update)
DE Adenylosuccinate lyase.
GN Name=purB; OrderedLocusNames=MW1849;
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwano N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RA "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
DR EMBL; AF004828; BAB95714.1; --
DR GO; GO:0016829; F:lyase activity; IEA.
DR InterPro; IPR003031; D:crystallin.
DR InterPro; IPR000362; Fumarate_lyase.
DR InterPro; IPR008948; L-Aspartase-like.

QY 33 GIFAIFSFYKLLQWNRKSKNKKRKEQIRBQIEGLLSYGAGVASLPLNVIHNPQS 92
DB 142 GVHAEPPTFGVKWALWYTEMQNLQRFQVREIEVGKMSGAVG-----TFANIPPE 193

QY 93 VISATPIYKGPCTGVPSNRLQITSGTAENTRILNHD 130
DB 194 IESYVCKHLG-----IGTAPVSTQTLQRD 217

RESULT 7
Q7A4Q3 PRELIMINARY; PRT; 431 AA.
AC Q7A4Q3
DT 05-JUL-2004 (TREMELrel. 27, Created)
DT 05-JUL-2004 (TREMELrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMELrel. 27, Last annotation update)
DE Adenylosuccinate lyase.
GN Name=purB; OrderedLocusNames=SA1724;
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita N., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RA "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003135; BAB42994.1; --
DR GO; GO:0016829; F:lyase activity; IEA.
DR InterPro; IPR003031; D:crystallin.
DR InterPro; IPR000362; Fumarate_lyase.
DR InterPro; IPR008948; L-Aspartase-like.
DR Pfam; PF00206; Lyase 1; 1.
DR PRINTS; PR00145; DCRYSTALLIN.
DR PRINTS; PR00149; FUMATELYASE.
DR TIGRPFAMs; TIGR00928; purB; 1.
DR Complete proteome; IEA.
KW Complete proteome; Lyase.
SQ SEQUENCE 431 AA; 49603 MW; 493F79CBE814B9E5 CRC64;

Query Match 11.7%; Score 84; DB 2; Length 431;
Best Local Similarity 26.5%; Pred. No. 12;
Matches 26; Conservative 15; Mismatches 35; Indels 22; Gaps 2;

QY 33 GIFAIFSFYKLLQWNRKSKNKKRKEQIRBQIEGLLSYGAGVASLPLNVIHNPQS 92
DB 142 GVHAEPPTFGVKWALWYTEMQNLQRFQVREIEVGKMSGAVG-----TFANIPPE 193

QY 93 VISATPIYKGPCTGVPSNRLQITSGTAENTRILNHD 130
DB 194 IESYVCKHLG-----IGTAPVSTQTLQRD 217

```

```
RESULT 8
QBDGCI
ID Q8DGC1 PRELIMINARY; PRT; 574 AA.
AC Q8DGC1;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE ABC transporter permease protein.
GN OrderedLocusNames=tl12397;
OS Synecococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=32046;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BP-1;
MEDLINE=2225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsumoto A., Nakazaki N.,
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RA "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1.";
RL DNA Res. 9:123-130 (2002).
CC -!- FUNCTION: Part of a binding-protein-dependent transport system.
CC Probably responsible for the translocation of the substrate across
CC the membrane (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the binding-protein-dependent transport
CC system permease family.
DR EMBL; AF005377; BAC09949.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; BPD_transp.
DR Pfam; PF00528; BPD_transp_1; 2.
DR PROSITE; PS50928; ABC_TM1; 2.
KW Complete proteome; Transmembrane; Transport.
SQ SEQUENCE 574 AA; 64027 MW; 3A478A3F80FB4BBD CRC64;

Query Match 11.7%; Score 84; DB 2; Length 574;
Best Local Similarity 27.6%; Pred. No. 17;
Matches 24; Conservative 20; Mismatches 27; Indels 16; Gaps 3;

QY 34 IPATIFPLFKLLOWNRKKNKRRQREIQELGLISYGAGVASLPILNIAHNPQSV 93
DB 74 MISLVFSILYSIAYNRTA-----EKILPLLDI---LQSIPLVSLF---PGVV 117

QY 94 ISATPIYKPGCTGVNPSRLQITSGTA 120
DB 118 LALIALFPGRNGVELAAIILYTGWA 144

RESULT 9
PURA AQUAE
ID PUR6 AQUAE STANDARD; PRT; 167 AA.
AC O67239;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Phosphoribosylaminoimidazole carboxylase catalytic subunit
DE (EC 4.1.1.21) (AIR carboxylase) (AIRC).
GN Name=purE; OrderedLocusNames=AQ_1178;
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=VF5;
MEDLINE=98196666; PubMed=9537320; DOI=10.1038/32831;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.B., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olsen G.J., Swanson R.V.;

RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RL Nature 392:353-358 (1998).
CC -!- FUNCTION: This subunit can alone transform AIR to CAIR, but in
CC association with purK, which possesses an ATPase activity, an
CC enzyme complex is produced which is capable of converting AIR to
CC CAIR efficiently under physiological condition (By similarity).
CC -!- CATALYTIC ACTIVITY: 5-amino-1-(5-phospho-D-ribose)imidazole-4-
CC carboxylate + 5-amino-1-(5-phospho-D-ribose)imidazole + CO(2).
CC -!- PATHWAY: De novo purine biosynthesis; sixth step.
CC -!- SUBUNIT: Homocotamer (By similarity).
CC -!- SIMILARITY: TO OTHER BACTERIAL PURE, ALSO TO AIR CARBOXYLASE FROM
CC FUNGI.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AE000727; AAC07201.1; -.
DR HSP; D70401; D70401.
DR HSP; Q9WYS7; I04V.
DR InterPro; IPR000031; AIR_carboxyl.
DR Pfam; PF00731; AIRC; 1.
DR ProDom; PD002193; AIR_carboxyl; 1.
DR TIGRFam; TIGR01162; PurE; 1.
KW Complete proteome; Decarboxylase; Lyase; Purine biosynthesis.
SQ SEQUENCE 167 AA; 18037 MW; 99A62BE600F50F4 CRC64;

Query Match 11.5%; Score 83; DB 1; Length 167;
Best Local Similarity 33.8%; Pred. No. 5;
Matches 23; Conservative 15; Mismatches 28; Indels 2; Gaps 1;

QY 5 HSIQVPPAPPNAPFVEIVSSSTGIIIVGIFAFISFLYKLLQWNRKSKNK--KKKEQI 62
DB 100 YSIVQMPAGIPVATVIGNATNAGLLAVRLISIKYPEYAKKLDEYTKLKEKVAQNEEL 159

QY 63 REQTELGL 70
DB 160 QKEVENGI 167

RESULT 10
Q22040 PRELIMINARY; PRT; 636 AA.
ID Q22040
AC Q22040;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Hypothetical protein SSSD1.1.
GN ORFNames=SSSD1.1;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Bristol N2;
MEDLINE=99069613; PubMed=9851916;
RA Wilson R.;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Bristol N2;
RA Minx M.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
```



```
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; U41530; AAA83273.3; -.
DR PIR; T16774; T16774.
DR HSP; P56276; 1FHG.
DR WormPep; SSSD1.1; CE29458.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR Pfam; PF00441; fn3; 1.
DR Pfam; PF00447; ig; 3.
DR SMART; SM00060; FN3; 1.
DR SMART; SM00409; IG; 3.
DR SMART; SM00408; IGC2; 3.
DR PROSITE; PS0853; FN3; 1.
DR PROSITE; PS0835; IG-LIKE; 3.
DR Hypothetical protein.
SQ SEQUENCE 636 AA; 71728 MW; 2894A1E7428C0793 CRC64;

Query Match 11.5%; Score 82.5; DB 2; Length 636;
Best Local Similarity 23.7%; Pred. No. 27;
Matches 32; Conservative 19; Mismatches 43; Indels 41; Gaps 7;

QY 10 PPAPPNAFVEIVSSSTG---IIIVGIFAFIFSFYKLLQW---YNRKSKNKKRKEQ--- 61
Db 469 PPSFNSALRYGIADNVLPFIPMLAAVFLGFCFACGLFAWRCHFNKKNKSKRKSRSNS 528

QY 62 ---IRQIEIGLUSYGAGVASLPLNLVIAHPNPGSVISATPIYKGPCTGVNSRLQITSGT 119
Db 529 TPSTKYQDYGRFYGDSSS-----SQPG-----TETY-----EPSRLLD----- 565

QY 120 AEENTRLNHDGRNP 134
Db 566 -----DHEWRGP 572

RESULT 11
QY Q70696 PRELIMINARY; PRT; 734 AA.
AC Q70696;
DT 01-MAR-2004 (TREMBlrel. 26, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DE AGCF5939 (Fragment)
GN Name=agCG54493; ORFName=ENSGANG00000008214;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (Mar-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAA01008960; EAA11826.1; -.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007092; LRR_SDS22.
DR InterPro; IPR003591; LRR_cyp.
DR Pfam; PF00560; LRR; 10.
DR PRINTS; PRO0019; LEURICHRPT.
FT NON_TER 1 734
FT NON_TER 734 734
SQ SEQUENCE 734 AA; 80917 MW; A409AFD311EAB054 CRC64;

STRAIN=Bristol N2;
Waterston R.;
Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
[4]
SEQUENCE FROM N.A.
STRAIN=Bristol N2;
Wilson R.;
Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
EMBL; U41530; AAA83273.3; -.
PIR; T16774; T16774.
HSP; P56276; 1FHG.
WormPep; SSSD1.1; CE29458.
InterPro; IPR003961; FN_III.
InterPro; IPR003599; IG.
InterPro; IPR007110; IG-like.
InterPro; IPR003598; IG_c2.
Pfam; PF00441; fn3; 1.
Pfam; PF00447; ig; 3.
SMART; SM00060; FN3; 1.
SMART; SM00409; IG; 3.
SMART; SM00408; IGC2; 3.
PROSITE; PS0853; FN3; 1.
PROSITE; PS0835; IG-LIKE; 3.
Hypothetical protein.
SEQUENCE 636 AA; 71728 MW; 2894A1E7428C0793 CRC64;

Query Match 11.4%; Score 82; DB 2; Length 734;
Best Local Similarity 26.9%; Pred. No. 35;
Matches 35; Conservative 15; Mismatches 36; Indels 44; Gaps 6;

QY 8 VOPPAPPNAFVEIVSSST-----GIIVGIFAFIFSFYKLLQWYNRKSKNKK 57
Db 620 VAPPPPPPPAAAVPVPSAAPLAAPILPFGVPLALATVAPI-PILYE----- 665

QY 58 RKEQIREQIEIGLUSYGAGVASLPLNLVIAHPNPGSVISATPIYKGPCTGVNSRL 112
Db 666 -----ASTSAPPPPPPVVVAHPHMGSGVSVVAASPTNGPLSSLENT-- 710

QY 113 LQITSGTAAE 122
Db 711 IQKTHATDDE 720

RESULT 12
QY Q8CRT6 PRELIMINARY; PRT; 431 AA.
AC Q8CRT6;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Adenylosuccinate lyase.
GN OrderedLocusNames=SEI593;
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12228;
RX PubMed=12950922;
RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J.,
RA Qiu Z.-Q., Miao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,
RA Yuan Z.-H., Zhao G.-B., Qu D., Danchin A., Wen Y.-M.;
RT "Genome-based analysis of virulence genes in a non-biofilm-forming
RT Staphylococcus epidermidis strain (ATCC 12228).";
RL Mol. Microbiol. 49:1577-1593(2003).
DR EMBL; AB016749; AAC05192.1; -.
DR HSP; G9X010; IC3C.
DR GO; GO:0004018; F:adenylosuccinate lyase activity; IEA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0009152; P:purine ribonucleotide biosynthesis; IEA.
DR InterPro; IPR003031; D_cristallin.
DR InterPro; IPR000362; Fumarate_lyase.
DR InterPro; IPR008948; L-Aspartase-like.
DR Pfam; PF00206; Lyase 1; 1.
DR PRINTS; PRO0145; DCRYSTALLIN.
DR PRINTS; PRO0149; FUMARATELYASE.
DR TIGRFAMs; TIGR00928; purB; 1.
DR PROSITE; PS00163; FUMARATE_LYASES; 1.
KW Complete proteome; Lyase.
SQ SEQUENCE 431 AA; 49591 MW; 08D770CAB109B72A CRC64;

Query Match 11.2%; Score 80.5; DB 2; Length 431;
Best Local Similarity 34.0%; Pred. No. 27;
Matches 17; Conservative 13; Mismatches 19; Indels 1; Gaps 1;

QY 33 GIFAFIFSFYKLLQWYNRKSKNKKRKEQIREQIEIGLUSYGAGV-ASLP 81
Db 142 GVHAEPTTFGVNMAWYTEMKNLKRPFKEVRKEIEVGMSGAVGTANIP 191

RESULT 13
QY Q6F2A9 PRELIMINARY; PRT; 443 AA.
AC Q6F2A9;
DT 01-OCT-2004 (TREMBlrel. 28, Created)
DT 01-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 01-OCT-2004 (TREMBlrel. 28, Last annotation update)
```



```
Db 579 -ILSYA--ASL-YTSLIYN-----MTTPVYTLKGKATQISSPFLDDSSGDEDS 629
Qy 123 --NTRILNHDGRNPDG 136
Db 630 RSSRLSESDARSRG 645

RESULT 15
Q8LP72 PRELIMINARY; PRT; 633 AA.
AC Q8LP72;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Receptor-like protein kinase.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamiales; Solanales; Solanaceae; Nicotiana.
QX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2191716; PubMed=11917080;
RA Ito N., Takabatake R., Seo S., Hiraga S., Mitsuhashi I., Ohashi Y.;
RT "Induced expression of a temperature-sensitive leucine-rich repeat
RT receptor-like protein kinase gene by hypersensitive cell death and
RT wounding in tobacco plant carrying the N resistance gene.";
RL Plant Cell Physiol. 43:266-274(2002).
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AB073628; BAC07504.2; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0006740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007090; LRR_plant.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00560; LRR; 6.
DR Pfam; PF00069; Kinase; 1.
DR PRINTS; PR00019; LEURICHRPT.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Receptor; Serine/threonine-protein kinase;
KW Transferase.
SQ SEQUENCE 633 AA; 70540 MW; 9221FF79372B4CD8 CRC64;
```

```
Query Match 11.0%; Score 79; DB 2; Length 633;
Best Local Similarity 23.7%; Pred. No. 59;
Matches 40; Conservative 18; Mismatches 51; Indels 60; Gaps 6;

Qy 11 PAPPNAFVEIV-----SSSTGIITAVGIFAFIFSFYKLLQWYR 51
D5 255 PAPASHNVAPVHKKKKRVKRVNPLGFGFAGSAV-----LLSVLFKLVNFFVR 309
Qy 52 KSRNKKR----KEIQEIQELGLSYGAGVASPLNLNVIAHNPFGSVISATPIYKGFCTGV 107
Db 310 KGKTDGSLTIYSPLIKKAEDLAFLEKEDGVASLEMG-----KGGCGEV 353
Qy 108 -----PNSRLLOI-----TSGTAENTRILNHDGRNPDGGINV 140
Db 354 YRAELPGSNGKIIAIKKILQPPMDAAELAEEDTKALNKKQKRVKSEIQI 402
```

Search completed: November 15, 2004, 14:57:34  
Job time : 198 secs